Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE ANALYST.

VOL. VII. NOVEMBER, 1880. No. 6.

ON A THEOREM IN PROBABILITY.

BY E. L. DE FOREST, WATERTOWN, CONN.

IT should be noticed in connection with the first property discussed in my paper on polynomials (ANALYST, March, 1880), that component adjustment formulas, such as (4) and (11) for example, need not satisfy the condition mentioned at page 41, of having the sum of the coefficients l or L equal to unity, unless the formulas are to be used for adjustments or other special purposes. The l and the L may have any values whatever, and the application of the two components will still have the same effect as that of one resultant formula with coefficients which are those of the product of the two polynomials corresponding to the components. So too, if any number of components are successively used, and in any order, they are together equivalent to a single resultant whose coefficients are those of the continued product of the corresponding polynomials.

The whole property seems to be related to a theorem in probability which was first given in its most general form by Lagrange. (See the complete edition of his works, Paris, 1868, Vol. II, p. 198.) With some changes of notation and method, it may be stated thus. Suppose that observations of any class are subject to given errors, or deviations from the true value of the observed quantity, and that each of these errors may occur in a given number of ways or cases, and that all the ways are equally probable. The errors a, b, c, &c., may be essentially either positive or negative, and they can occur in α, β, γ, &c., ways respectively. Then in k observations, the number of ways or cases in which the algebraic sum of the errors can be v, is the coefficient of x^v in the expansion of the polynomial

$$(\alpha x^a + \beta x^b + \gamma x^c + \&c.)^k, \tag{1}$$

The probability that the sum will be v is therefore the coefficient of x^v in the expansion of

$$\left(\frac{\alpha x^a + \beta x^b + \gamma x^c + \&c.}{\alpha + \beta + \gamma + \&c.}\right)^k, \tag{2}$$

which may be written

$$(px^a + qx^b + rx^c + \&c.)^k, \tag{3}$$

where p, q, r, &c., are the probabilities of the occurrence of the errors a, b, c, &c., respectively, in a single observation, and

$$p + q + r + \&c. = 1. \tag{4}$$

The same coefficient of x^v in the expansion of (3) is also the probability that the error of the mean of the k observations is $v \div k$.

The most probable total effect of the error a, in k observations, is kpa, that of b is kqb, that of c krc, and so on, wherefore the most probable algebraic sum of the errors is

$$k(pa + qb + rc + \&c.), \tag{5}$$

and this is approximately the exponent of that power of x whose coefficient in the expansion of (3) is a maximum. The most probable error in the mean of the k observations is therefore

$$pa + qb + rc + \&c. \tag{6}$$

The above is Lagrange's theorem in substance. An abstract of it may be found in Todhunter's *History of the Theory of Probability*, p. 304. We now observe further, that the exponents a, b, c, &c., which represent the errors, can be regarded as whole numbers, for if they were fractions, they might be reduced to whole numbers by multiplying them all by a common multiple of their denominators, which amounts simply to measuring the errors on a reduced scale. And the exponents may all be made positive, for if the largest negative one is $-h$ for example, we can add h to each of them, and so convert (3) into the form

$$(p_0 + p_1 z + p_2 z^2 + \ldots + p_m z^m)^k, \tag{7}$$

where z^0 and z^{v+h} take the places of x^{-h} and x^v respectively. Each term in the first power of (3) is now represented by some term in the first power of (7). Those terms in the first power of (7) which do not represent any term in the first power of (3), will of course have their coefficients p equal to zero, so that

$$p_0 + p_1 + p_2 + \ldots + p_m = 1. \tag{8}$$

The coefficients in the expansion of (3) to the k power are not altered by multiplying all the exponents of the first power of the polynomial by a constant number, or by adding a constant number to all of them, or by both these changes together. Hence it follows that the properties of the lever arm and the radius of gyration of the coefficients in powers and products of polynomials, demonstrated by me in the ANALYST, March and May, 1880, pp. 44 and 74, will hold good alike whether the exponents are positive or negative, integral or fractional, provided that we locate the coefficients along the axis or lever at distances from a given point which are equal or proportional to the corresponding exponents. The lever arm of the coefficients in the product of two or more polynomials, about the place of exponent zero, is equal to the sum of the lever arms of all the factors, about their respective zero places. We represent by Δx that interval on the lever arm which corresponds to a difference of unity in the exponents of (7).

It may be well to prove the above directly for the two polynomials

$$\left.\begin{array}{l} \alpha_{-m}z^{-a_m}+\ldots+\alpha_{-2}z^{-a_2}+\alpha_{-1}z^{-a_1}+\alpha_0z^0+\alpha_1z^{b_1}+\alpha_2z^{b_2}+\ldots+\alpha_mz^{b_m}, \\ \beta_{-n}z^{-c_n}+\ldots+\beta_{-2}z^{-c_2}+\beta_{-1}z^{-c_1}+\beta_0z^0+\beta_1z^{d_1}+\beta_2z^{d_2}+\ldots+\beta_nz^{d_n}, \end{array}\right\} \qquad (9)$$

which have both positive and negative exponents. Let the sum of the coefficients α in the first polynomial be S_1, and let their lever arm about the place of α_0 as a fulcrum be h_1. Also let the sum of the coefficients β in the second polynomial be S_2, and let their lever arm about the place of β_0 be h_2. The statical moments of the two systems about these fulcrums are

$$\left.\begin{array}{l} S_1h_1=(-\alpha_{-m}a_m-\alpha_{-m+1}a_{m-1}-\ldots-\alpha_{-1}a_1+\alpha_1b_1+\alpha_2b_2+\ldots+\alpha_mb_m)\Delta x, \\ S_2h_2=(-\beta_{-n}c_n-\beta_{-n+1}c_{n-1}-\ldots-\beta_{-1}c_1+\beta_1d_1+\beta_2d_2+\ldots+\beta_nd_n)\Delta x. \end{array}\right\} \qquad (10)$$

Now when the first one of the polynomials (9) is multiplied by the second, the moments about the place of $\alpha_0\beta_0$, of the coefficients in the parts of the product due to $\beta_0, \beta_1, \beta_2, \ldots \beta_n$, are

$$\beta_0S_1h_1,\ \beta_1S_1(h_1+d_1\Delta x),\ \beta_2S_1(h_1+d_2\Delta x), \ldots \beta_nS_1(h_1+d_n\Delta x),$$

and those for the parts due to $\beta_{-1}, \beta_{-2}, \ldots \beta_{-n}$ are

$$\beta_{-1}S_1(h_1-c_1\Delta x),\ \beta_{-2}S_1(h_1-c_2\Delta x), \ldots \beta_{-n}S_1(h_1-c_n\Delta x).$$

The sum of all the coefficients in the product is S_1S_2, and denoting by H their lever arm about the place of $\alpha_0\beta_0$, we have, by addition of the foregoing moments,

$$\begin{aligned} S_1S_2H = & \ S_1h_1(\beta_{-n}+\beta_{-n+1}+\ldots+\beta_{-1}+\beta_0+\beta_1+\beta_2+\ldots+\beta_n) \\ & + S_1(-\beta_{-n}c_n-\beta_{-n+1}c_{n-1}-\ldots-\beta_{-1}c_1+\beta_1d_1+\beta_2d_2+\ldots+\beta_nd_n)\Delta x, \end{aligned}$$

which by the help of (10) is reduced to

$$S_1S_2H = S_1S_2h_1 + S_1S_2h_2,$$

$$\therefore H = h_1 + h_2, \qquad (11)$$

so that the lever arm in the product is equal to the sum of the lever arms in the two factors.

We will now make another statement and application of Lagrange's theorem. An urn contains a number of balls, some of which are marked 0, others 1, others 2, &c., in such proportion that the probabilities of drawing a single ball marked 0, 1, 2, ... m are $p_0, p_1, p_2, \ldots p_m$, respectively.

The occurrence of an error b, under the first statement, is here represented by the drawing of a ball marked $b+h$. If each ball is returned as soon as drawn, and they are all mixed so as to keep the chances the same, then after k drawings, the probability that the sum of the numbers drawn shall be a given number s, is the coefficient of z^s in the expansion of (7). Proof of this will be given farther on. The most probable shares which the numbers 0, 1, 2, &c., will respectively have in the formation of the sum, are

$$kp_0 \times 0, \quad kp_1 \times 1, \quad kp_2 \times 2, \text{ \&c.}, \ kp_m \times m,$$

so that the most probable sum will be

$$k(p_1 + 2p_2 + 3p_3 + \ldots + mp_m). \qquad (12)$$

This is the value, approximately, of the exponent of z in that term of the expansion whose coefficient is a maximum. It is in general approximate only, because it is not usually a whole number, as the exponents here are supposed to be. It will be seen that the portion of (12) within the parentheses is the lever arm of the coefficients p in the first power of (7), about the place of p_0 as a fulcrum, the constant interval Δx between the coefficients being unity. This answers to the property noticed by Lagrange in his first corollary. I have shown further, that the lever arm of the coefficients in the expansion to the k power will be k times what it is in the first power. Hence according to (12), the maximum coefficient in the expansion will be located approximately at the centre of forces, or centre of gravity, of the whole expanded series of coefficients, even when k is not a large number. I have before shown that this is true at the limit, when k becomes large or infinite, and proved that the series of coefficients then takes the form of the probability curve. (ANALYST, Sept. and Nov., 1879.) For a given value of k, the place or rank of the vertex is known and the curve can be constructed. Since k in my article last mentioned is in strictness infinite, we may write k instead of $k+1$. Also $b_2(dx)^2$ there represents the square of the radius of gyration r, of the coefficients of the first power of the polynomial, about their centre of forces. The equation of the curve then is

$$h^2 = \frac{1}{2kr^2}, \qquad y = \frac{hdx}{\sqrt{\pi}} e^{-h^2x^2}, \qquad (13)$$

where dx is what the interval Δx reduces to when the coefficients, regarded as ordinates, are brought together so as to be consecutive, and kr^2 is the square of the radius of gyration of the coefficients in the expanded polynomial, about their centre of forces. This radius is here seen to be identical with the quadratic mean error; as we have already noticed in the case of the expansion of the binomial $(p+q)^m$. (ANALYST, Jan., 1880, pp. 8 and 22.) The probable error, or probable deviation of the sum of the numbers drawn, from its most probable value, is

$$.6745\left(\frac{r}{dx}\right)\surd k. \tag{14}$$

In a special case, where the coefficients p are all equal, Laplace found, by a very different method, that the limiting form of the expansion is what we call the probability curve. (*Théorie analytique des Probabilités*, edition of 1847, pp. 161 and 333.) He took this as a basis for demonstrating the exponential law of probability, and extended it to the case of given errors distributed according to any algebraical law, provided it be such that positive and negative errors of equal amount are equally probable. An account of the method may be found in Airy's *Theory of Errors of Observations*, pp. 7 to 15. De Moivre had first shown that the probability curve is the limiting form of the expansion of $(\frac{1}{2}+\frac{1}{2})^m$. (Todhunter, p. 192.) The special case above named where the coeff. p are all equal, arises when when we make a number of throws with dice, and wish to find the probability the sum of the numbers thrown will be a given number, or, fall within given limits.

The general theorem of Lagrange can be extended to include the case of n urns, which we will suppose to contain balls bearing either positive or negative numbers, in such proportions that the probabilities of drawing the several numbers from $-m$ to $+m$ are $p'_{-m}, \ldots p'_0, p'_1, p'_2, \ldots . p'_m$ for the first urn, $p''_{-m}, \ldots p''_0, p''_1, p''_2, \ldots p''_m$ for the second, etc., and $p^{(n)}_{-m}, \ldots p^{(n)}_0, p^{(n)}_1, p^{(n)}_2, \ldots p^{(n)}_m$ for the nth urn, m being the highest number, positive or negative, on any ball. Then if we draw one ball from each urn, the probability that the sum of all the numbers drawn shall be s, will be the coefficient of z^s in the continued product of the polynomials

$$\left.\begin{array}{l} p'_{-m}z^{-m}+\ \ldots\ +p'_0+p'_1z+p'_2z^2+\ \ldots\ +p'_mz^m, \\ p''_{-m}z^{-m}+\ \ldots\ +p''_0+p''_1z+p''_2z^2+\ \ldots\ +p''_mz^m, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p^{(n)}_{-m}z^{-m}+\ \ldots\ +p^{(n)}_0+p^{(n)}_1z+p^{(n)}_2z^2+\ldots+p^{(n)}_mz^m. \end{array}\right\} \tag{15}$$

For this coefficient of z^s is the sum of all the products of n factors p taken in the order $p'p''p''' \ldots p^{(n)}$, which can be formed in such manner that the sum of all the sub-indices in each product shall be s; and each of these products is the probability of drawing, in that particular order, the numbers denoted by the sub-indices. If we draw k balls from a single urn, replacing each as before, we must merely suppose that k polynomials have like coefficients. Hence in general, if we make k' drawings from the first urn, k'' from the second, and so on to the nth, the probability that the sum of all the numbers drawn shall be s, will be the coefficient of z^s in the continued product of the expansions of the polynomials

$$\left.\begin{array}{l} (p'_{-m}z^{-m}+\ \ldots\ +p'_0+p'_1z+p'_2z^2+\ \ldots\ +p'_mz^m)^{k'}\ , \\ (p''_{-m}z^{-m}+\ \ldots\ +p''_0+p''_1z+p''_2z^2+\ \ldots\ +p''_mz^m)^{k''}\ , \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (p^{(n)}_{-m}z^{-m}+\ \ldots . p^{(n)}_0+p^{(n)}_1z+p^{(n)}_2z^2+\ \ldots\ +p^{(n)}_mz^m)^{k(n)}. \end{array}\right\} \tag{16}$$

When the exponents k are tolerably large, there is no need to actually perform the multiplications. My theorems respecting the lever arm and the radius of gyration of the coefficients enable us easily to find the centre of forces for the final product, giving the most probable value of s, and to find the radius of gyration about this centre, determining h in the probability curve (13). The proposition that the most probable value is at the centre of forces approximately, is also in accordance with the known theorem in mechanics, that the radius of gyration (here the mean error) is a minimum when the axis of rotation is taken through the centre of gravity.

To illustrate the foregoing, suppose that two urns contain ten balls each. The first has four balls marked —1, five marked 0, and one marked 2; while the second has one marked —2, two marked —1, four marked 1, and three marked 2. Twelve drawings are made from the first and eight from the second. The two polynomials are

$$\left(\frac{4z^{-1}+5+z^2}{10}\right)^{12}, \qquad \left(\frac{z^{-2}+2z^{-1}+4z+3z^2}{10}\right)^{8}, \tag{17}$$

and the lever arms of the coefficients in their first powers, about the places of z^0 as fulcrum, are

$$\left.\begin{aligned} \tfrac{1}{10}[4(-1)+5(0)+1(2)]\Delta x &= -\tfrac{1}{5}\Delta x, \\ \tfrac{1}{10}[1(-2)+2(-1)+4(1)+3(2)]\Delta x &= \tfrac{3}{5}\,\Delta x. \end{aligned}\right\} \tag{18}$$

Hence, in the product of the expansions of (17), the lever arm, about the place of z^0 as a fulcrum, is

$$12(-\tfrac{1}{5}\Delta x)+8(\tfrac{3}{5}\Delta x) = \tfrac{12}{5}\Delta x. \tag{19}$$

The integer nearest to $\frac{12}{5}$ is 2, which is therefore the most probable algebraic sum of the numbers drawn. The squares of the radii of gyration of the coefficients in the first powers of the polynomials (17), about the places of the extremity of the lever arm, or the centre of parallel forces, are

$$\left.\begin{aligned} \tfrac{1}{10}[4(\tfrac{4}{5})^2+5(\tfrac{1}{5})^2+1(\tfrac{11}{5})^2]\,(\Delta x)^2 &= \tfrac{19}{25}(\Delta x)^2, \\ \tfrac{1}{10}[1(\tfrac{13}{5})^2+2(\tfrac{8}{5})^2+4(\tfrac{2}{5})^2+3(\tfrac{7}{5})^2]\,(\Delta x)^2 &= \tfrac{46}{25}(\Delta x)^2. \end{aligned}\right\} \tag{20}$$

Therefore in the product of the expansions of (17), the square of the radius of gyration of the coefficients, about their centre of forces, is

$$[12(\tfrac{19}{25})+8(\tfrac{46}{25})]\,(\Delta x)^2 = 23.84(\Delta x)^2. \tag{21}$$

The probable error then is

$$.6745\sqrt{(23.84)} = 3.29,$$

so that it is about an even chance that the sum of the numbers drawn will fall within the limits $\frac{12}{5} \pm 3.29$.

The product of the expansions of (17) will approximate to the form of a probability curve. I have already shown that this is true for any high power of a single polynomial whose coefficients are all positive, so that the two expansions in (17) will each be of that form. Now in the curve (13) the parameter h is constant so long as kr^2 is constant, though k and r^2 should vary inversely as each other; and many different polynomials may give the same value of r^2. Hence a curve of given parameter may be regarded as approximately representing the form of the expansion, to some power, of an indefinite number of different polynomials. The two expansions in (17) may thus be taken as expansions of some one undetermined polynomial, to different powers. Their product will thus be an expansion of that polynomial, and consequently will take the form of the probability curve.

To construct its equation from (13), putting dx for Δx in the radius of gyration, we have

$$h^2 = \frac{1}{2\times 23.84(dx)^2}, \qquad \therefore hdx = .14482,$$

and writing $i = x \div dx$, we get

$$\left.\begin{aligned} \log y &= \log(hdx) - \tfrac{1}{2}\log\pi - i^2(hdx)^2\log e \\ &= \bar{2}.91226 - .0091085\,i^2. \end{aligned}\right\} \qquad (22)$$

The vertex of the curve, where $i = 0$, is at the centre of forces in the product of the expansions of (17), and the rank of that centre from z^0, as given in (19), is $\frac{12}{5}$. Hence the place of the exponent 2 corresponds to $i = -\frac{2}{5}$, so that if we assign to i the values $-\frac{7}{5}$, $-\frac{2}{5}$, $\frac{3}{5}$, $\frac{8}{5}$, &c., the resulting values of y will be the approximate probabilities that the sum of the numbers drawn will be 1, 2, 3, 4, &c., respectively. The larger the exponents k are in such cases, the more accurate will the results be.

To show the general correctness of the above approximation, the computed values of y, for each value of i, are shown in the subjoined table. In the column headed C are also placed the values of the coefficients in the product of the expansions (17), as found by actual multiplication.

i	y	C	i	y	C	i	y	C	i	y	C
—15.4	.001		—7.4	.026	.027	0.6	.081	.083	8.6	.017	.018
—14.4	.001		—6.4	.035	.035	1.6	.077	.077	9.6	.012	.011
—13.4	.002		—5.4	.044	.045	2.6	.071	.072	10.6	.008	.006
—12.4	.003	.004	—4.4	.054	.054	3.6	.062	.061	11.6	.005	.005
—11.4	.005	.005	—3.4	.064	.065	4.6	.052	.052	12.6	.003	
—10.4	.008	.007	—2.4	.072	.074	5.6	.042	.042	13.6	.002	
— 9.4	.013	.013	—1.4	.078	.078	6.6	.033	.033	14.6	.001	
— 8.4	.019	.019	—0.4	.081	.083	7.6	.024	.023	15.6		

Only three places of decimals were used in doing this part of the work, so that the series C is not perfectly smooth and regular, but its general agreement with the series y is very close.

The probabilities that an event will happen or fail certain numbers of times in m trials, are represented, as is well known, by terms in the expansion of $(p+q)^m$. (ANALYST, May 1879, p. 65.) This is the simplest case under our present theorem, where there is but a single urn, containing balls marked either 0 or 1 for example, so that (16) reduces to

$$(p'_0+p'_1z)^{k'}. \tag{23}$$

The coefficient of z^s in the expansion is the probability that the event 1 will happen s times, and consequently that the event 0 will happen $k'-s$ times.

Reverting now to our first statement of Lagrange's theorem, as applied to the probable effect of a system of given errors a, b, c, &c., it appears that the most probable error in the mean result of k observations is the kth part of that total error, or algebraic sum of the errors, whose probability occupies the centre of forces in the series of coefficients in the expansion of the polynomial (3), when the coefficients in this expansion are regarded as masses acting about the point of no error as a fulcrum, with lever arms equal to the total errors whose probabilities are represented by the coefficients. The probable sum of the total deviations from the centre of forces or centre of gravity on one side, is equal to the probable sum of those on the other, so that the most probable mean result of the k observations is the arithmetical mean of all the mean results possible under the given system of errors, each possible result having a weight proportional to the prob'y of its occurrence.

If the algebraic sum of the given elementary errors, each weighted in proportion to the probability of its occurrence, is zero, the most probable mean result will be the true value of the observed quantity; and not otherwise. The most probable result will be greater or less than the true value according as the + or — elementary errors preponderate. When these errors are unknown, we have no reason to expect a preponderance on one side rather than on the other, and hence we conclude that as far as the true value of an observed quantity can be inferred from the results of numerous observations, the arithmetical mean of all the observed values is the most likely to be the true value. It is to be presumed that the observed deviations from the mean will be so distributed as to represent approximately the distribution of all the possible deviations.

As the number of observations is increased, the series of coeff's in the expansion of (3) approaches rapidly, as I have shown, to the form of the probability curve (13), and attains this form at the limit, when k becomes very large, or infinite. It appears to me that we have here the most general and complete proof of the validity of that exponential law of facility of error, or probability of deviation from the most probable value, whose geometrical expression is called the probability curve. The elementary errors are not restricted here by assuming that they follow an algebraical law of distribution, nor are positive and negative errors of equal amount required to be equally probable.

I will take a future occasion to show the application of similar principles in determining the law of errors in two dimensions, or errors in the position of points in a plane.